AD_____

Award Number: W81XWH-11-1-0549

TITLE: Enhancing the Breadth and Efficacy of Therapeutic Vaccines for Breast Cancer

PRINCIPAL INVESTIGATOR: Paul T. Spellman, PhD

CONTRACTING ORGANIZATION: Oregon Health & Science University
Portland, OR 97201

REPORT DATE: October 2014

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved

OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2014	Annual	25 Sep 2013 – 24 Sep 2014

4. TITLE AND SUBTITLE

Enhancing the Breadth and Efficacy of Therapeutic Vaccines for Breast Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER

W81XWH-11-1-0549

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Paul Spellman, PhD; Dmitri Rosanov, PhD; Kami Chiotti, MS

E-Mail: spellmap@ohsu.edu, rosanov@ohsu.edu, chiotti@ohsu.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Oregon Health & Science University
Portland, OR 97201

8. PERFORMING ORGANIZATION REPORT

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The focus of the Spellman/Gray work group over the past year has been upon the generation of materials, tools, and data for the purpose of aiding and supporting the research and findings of the entire multi-team collaboration endeavoring to identify antigenic targets for breast cancer-infiltrating T cells. Our team has achieved a number of accomplishments over the current funding year. We have characterized immunogenic peptides from a collection of MHC-I-bound epitopes eluted from the cell surface of several breast cancer cell lines. A computational pipeline was also developed to identify the sequence of the complete TCR heterodimer, working synergistically with data collected following single-cell emulsion RT-PCR. Additional modifications were made to our epitope discovery workflow to increase efficacy of transcript and neoantigen candidate prioritization for future research, and strides are being made in development of a personalized T cell-based protocol for identification of T cell-activating epitopes.

15. SUBJECT TERMS NOTHING LISTED

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	12	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

	Page
Introduction	4
Keywords	4
Overall Project Summary	4
Key Research Accomplishments	9
Conclusion	10
Publications, Abstracts, and Presentations	10
Inventions, Patents, and Licenses	10
Reportable Outcomes	10
Other Achievements	10
References	10
Appendices	12

INTRODUCTION:

The OHSU Spellman/Gray work group is one of three collaborators funded by this Department of Defense Breast Cancer Multi-Team Award; the other two being comprised of the Lee work group from City of Hope (formerly of Stanford Medicine Cancer Institute) and the Slansky/Kappler work group from University of Colorado Denver/National Jewish Health. The major objective of this endeavor is to develop novel strategies aimed at the enhancement of the protective effects of anti-tumor T cells *in vivo* in a patient-specific manner based on the hypothesis that partially protective anti-tumor T cells exist within TDLNs in most breast cancer patients. This will be accomplished by identifying the antigens anti-tumor T cells target in different breast cancer subtypes, potentially including antigens preferentially expressed by breast cancer stem cells. We will identify both MHC-I- and MHC-II-restricted antigens driving both CD8 and CD4 anti-tumor T cells *in vivo*, as CD4 T cells are needed to optimally sustain vaccine-elicited CD8 T cells *in vivo* [1]. Identified antigens will be categorized as to breast cancer subtype-specificity or shared status amongst subtypes, with the intention a patient could be matched with an optimal set of vaccine antigens for her tumor. Another novel aspect of this project is the identification of altered peptides (mimotopes) that may more efficiently activate anti-tumor T cells than the natural tumor epitopes. A final objective is to identify small molecule anti-cancer agents that synergize with cytotoxic T lymphocytes (CTLs) to enhance immune-mediated killing. Collectively, this undertaking will produce a set of immunologically validated antigens and mimotopes for major breast cancer subtypes, and a set of agents that cooperate with immune killing. These can be used in combinations in a patient-specific manner to maximize clinical benefit while minimizing toxicity. The tools we develop will enhance the breadth and efficacy of existing and future approaches for immune therapy of breast cancer. We discuss here the Spellman/Gray group's specific efforts toward realizing the goals of this collaboration.

KEYWORDS:

Breast cancer, cytotoxic T lymphocytes, RNAseq, MiTCR, immune response, epitopes

OVERALL PROJECT SUMMARY:

Generation and initial analysis of T cell clones [Task 5]

Confirm tumor reactivity and HLA restriction of clones. The Spellman/Gray lab continues to contribute to the progress of this task by further interrogating the immunogenic HLA-A2-restricted epitopes eluted from the surface of breast carcinoma cells, which were reported in our 2013 annual report. The total numbers of eluted peptides reported as well as their corresponding proteins for each cell line are again provided in **Table 1**. The total numbers of peptides and their associated proteins, however, do not correspond to the numbers of unique peptides and unique proteins. This is due to MHC I-presented peptides and proteins that are shared amongst different breast carcinoma cell lines. Removing duplicate winnowed the total numbers of eluted peptides from 3358 to 2813 and the associated proteins from 3070 to 1939. Briefly we will review process by which we narrowed this list to identify those eluted epitopes most likely to have the ability to activate T cell response. We first selected genes with alterations in at least 20% of invasive breast

	Cell line	Subtype	№ peptides	FDR (%)	№ proteins	FDR (%)
1	SUM159PT	Claudin-low	**439**	13	**385**	13
2	MDA-MB-231	Claudin-low	**9**	10	**9**	9
2	MDA-MB-231	Claudin-low	**49**	15	**46**	15
2	MDA-MB-231	Claudin-low	**10**	6	**10**	20
3	HCC1395	Claudin-low	**83**	9	**81**	10
4	BT549	Claudin-low	**22**	1	**22**	20
5	HCC70	Basal	**271**	9	**251**	8
6	HCC1187	Basal	**688**	6	**607**	9
7	HCC1569	Basal	**200**	6	**189**	9
8	MCF12A	Basal	**87**	1	**83**	4
9	CAL-120	Basal	**4**	1	**4**	11
10	HCC1500	Basal	**33**	8	**32**	9
11	MDA-MB-468	Basal	**274**	6	**256**	7
12	HCC1806	Basal	**299**	6	**273**	9
13	LY2	Luminal	**241**	5	**226**	11
14	MCF7	Luminal	**222**	6	**203**	9
15	CAMA-1	Luminal	**118**	1	**104**	4
16	T47D HER2+	Luminal	**75**	1	**71**	9
17	HCC1419	Luminal	**17**	2	**17**	10
18	HCC1428	Luminal	**22**	1	**21**	7
19	SUM185PE	Luminal	**88**	2	**86**	2
20	UACC812	Luminal	**107**	2	**94**	3
	Total		**3358**		**3070**	
	Unique		**2813**		**1939**	

Table 1. Number of eluted MHC I-restricted peptides and corresponding proteins in breast carcinoma cells. FDR, false discovery rate.

cancers using data obtained from cBioPortal [2]. These alterations included copy number amplification, homozygous deletion, mRNA upregulation or downregulation, and mutation. Second, we used gene expression data from 708 breast tumors and 329 normal tissues (obtained from TCGA, EBI, and GEO [3–5]), 62 breast carcinoma cell lines, and 6 non-transformed cell lines to identify genes with preferential expression in breast cancer samples over normal (at least 4-fold difference). Finally, we selected epitopes and genes frequently identified by our MHC I immunoprecipitation and elution approach among the different cell lines (at least 4 times). The total number of epitopes meeting these criteria is 467. Our search was limited to HLA-A2-restricted epitopes, because the HLA-A2 allele is the most frequent allele amongst the US and Caucasian population. Using specialized software, the HLA-A2 binding scores were determined for each peptide, and peptides scoring less than 20 were filtered out. The highest scoring peptide is 36. Approximately 170 peptides were selected for further analysis.

During this funding year, we have determined which of the 170 peptides bind HLA-A2 molecules. For this purpose, we used T2 cells. T2 is a lymphoblastoid cell line with a mutated TAP gene and expresses HLA-A2 molecules without a loaded epitope; however, epitopes can be loaded exogenously. Loaded HLA-A2 molecules can be detected using HLA-A2-specific antibodies that recognized the loaded form of HLA-A2. As T2 cells are floating cells, we optimized conditions to attach T2 cells to a 96-well plate surface. We found T2 cells bind firmly to plates coated with concanavalin A (Con A). **Figure 1** shows the binding of positive control peptides (N181-183) to the HLA-A2 molecules expressed on the cell surface of T2 cells can be readily detected. We have determined approximately 120 of the selected peptides bind to HLA-A2 molecules on the T2 cell surface.

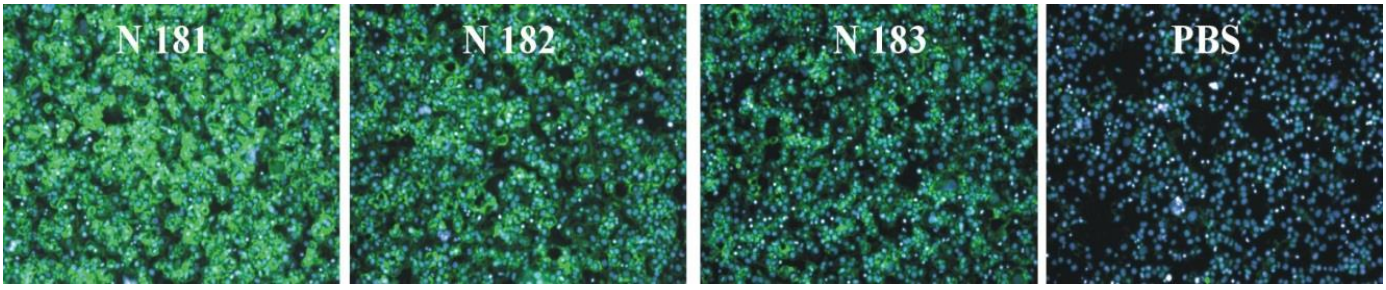

Figure 1. Staining of peptide-pulsed T2 cells with HLA-A2-specific antibodies.

To identify which selected peptides are actually immunogenic, we used a T cell activation protocol published by Wulfl et al [6] and Ho et al [7]. Dendritic cells (DCs) were generated from HLA-A2-positive peripheral blood mononuclear cells (PBMCs) with a 90-min incubation at 37°C in DC medium. Non-adherent cells and medium were removed and replaced with 2 mL/well fresh DC medium. The DC medium was supplemented with 1000 IU/mL GM-CSF and 1000 IU/mL IL-4. After one day of incubation, immature DCs were matured using 10 ng/mL lipopolysaccharide (LPS) and 50 IU/mL IFN-γ in the presence of peptide (10 µg/mL). The following day, the peptide-pulsed DCs were irradiated with 32 Gy, mixed with autologous CD8+ T cells, and incubated for seven days. On day four, IL-2 (50 IU/mL) and IL-7 (5 ng/mL) were added to the medium.

Secondary stimulation was carried out as described above with the exception HLA-A2-positive PBMCs were used and cytokines were supplemented after two days rather than four. After seven days incubation, cells were harvested and stained with CD137 antibodies to determine the ratio of activated T cells for each peptide. **Table 2** shows each peptide that induced CD137 expression on the T cell surface.

TCR sequencing of each clone. The CompleteClone pipeline was constructed to determine the repertoire diversity of T cell receptor clones from raw next generation sequencing data. CompleteClone is built on the foundation of the MiTCR [8] open source software package developed by MiLaboratory. MiTCR is a highly efficient and fast approach to CDR3 extraction, clonotype assembly, and repertoire diversity estimation while accounting for sequencing and PCR errors as well as salvaging low-quality input reads. Currently, MiTCR is limited to analysis of either the α chain or the β chain (human or mouse) of the TCR heterodimer. CompleteClone enhances the capabilities of MiTCR by allowing determination of TCR clone repertoire diversity of the matched αTCR-βTCR complex using the raw TCR sequence data of individual T cells generated by the Slansky work group. This is accomplished via downstream manipulation of MiTCR outputs using R [9].

Sequence	Symbol	Gene	%CD137+	Cell Line
ALQEASEAYL	H3F3A	H3 Histone, Family 3A	4.5	MCF7
LLQEVEHQL	TRIM37	Tripartite Motif Containing 371	5.6	MCF7
HLFEKELAGQSR	LAD1	Ladinin 1	6.8	HCC1187
LLDVPTAAV	IFI30	Interferon, Gamma-Inducible Protein 301	8.0	MDAMB231, SUM159PT, MCF7, LY2
LLGPRLVLA	TMED10	Transmembrane Emp24-Like Trafficking Protein 10 (Yeast)	6.2	MCF7
AGAMAGVMGAYL	SLC25A35	Solute Carrier Family 25, Member 35	6.8	SUM159PT
AAAGSPVFL	SLC16A3	Solute Carrier Family 16, Member 3 (Monocarboxylic Acid Transporter 4)	4.3	MDAMB231
FTEAGLKELSEY	BZW1	Basic Leucine Zipper and W2 Domain-Containing Protein 1	4.5	HCC1187
AEIDAHLVAL	PSMA6	Proteasome Subunit Alpha Type-6	5.5	HCC1187
ILTDITKGV	EEF2	Eukaryotic Translation Elongation Factor 2	5.8	HCC1500, MCF12A, SUM159PT, LY2, MCF7
SAQGSDVSLTA	HLA-B	Major Histocompatibility Complex, Class I, B	8.0	SUM159PT, HCC70
No Peptide			2.9	

Table 2. Peptides that induce CD137 expression.

Since MiTCR assigns each input read a numeric identifier, it was necessary to make two modest changes to the MiTCR source code in order to produce the output required by CompleteClone to match the α reads for each clonotype to their β mates. First, the standard MiTCR results file now includes a list of the numeric IDs for all reads belonging to each clonotype. Second, a temporary output file is created mapping the sequence identifier for each read in the input FASTQ file to its MiTCR-assigned numeric identifier. No changes were made to the algorithms MiTCR uses for CDR3 extraction, clonotype assembly, or error correction. The aforementioned R script first annotates the reads of each α clonotype with the appropriate sequence identifiers, repeating the process for the reads of each β clonotypes. The α and β reads are now paired by their sequence identifier, and any read lacking a mate is removed from the dataset. Finally, the frequencies of αTCR-βTCR pairs, or clonotypes, are calculated.

CompleteClone requires Java version 1.7.0 [10] or higher and R version 3.1.0 [9] or higher with the *plyr()* package [11]. It is run from the command line via a shell wrapper script that requires an input manifest detailing locations of the α and β FASTQ files as well as their corresponding sample names. This approach facilitates high throughput data processing.

RNAseq analysis of tumor cells [Task 7]

RNAseq analysis to identify breast cancer-specific aberrant transcripts. As previously reported, RNAseq datasets were used to conduct a systematic computational analysis to identify aberrant transcripts resulting in breast cancer antigens. In review, we developed an epitope prediction pipeline utilizing approximately 1000 breast cancer and normal tissue RNAseq samples available through TCGA, EBI, and GEO. The Tuxedo software suite [12–14] was used to carry out sequence assembly and alignment, prediction of novel isoforms, and quantitation of transcript structure. A collection of novel and known transcripts were predicted to be preferential to breast tumor tissue following Median Split Silhouette (MSS) clustering and a series of filtering steps. The filtered novel transcripts then individually underwent *in silico* validation to determine the exact peptide sequence differing from the reference genome.

Over the past year, we have made some adjustments to the epitope discovery pipeline, which are indicated by the red, bold-lined boxes in **Figure 2**. Namely, the modifications focused on 1) an enhanced method for ranking transcripts and epitopes as to expression specificity in tumor over normal tissues and 2) an automated workflow for discerning the unique portions of novel isoform sequences in large batches, rather than interrogating them individually.

Originally, we selected the strongest transcript candidates by setting arbitrary cutoffs on percentage of tumor population and expression level represented in a cluster of interest. Transcripts failing to meet the set criteria were discarded. As we are also interested in where known immunogenic transcripts fall within the dataset and this requiring the dataset remain intact, we established a heuristic equation for *expression ranking* (**Fig. 2**) to calculate the rank of every candidate exhibiting a bimodal (high and low) or trimodal (high, mid, and low)

expression profile across all samples. The equation is designed to highlight tumor-specific transcripts by placing higher weight on those where the high expression (H) cluster:

1) Is comprised predominantly of tumor samples as determined by the number of tumor samples (TS_H) present in the cluster population (CP_H):

   $$Tumor\ Fraction\ (TFx) = TS_H \diagup CP_H$$

2) Represents a significant portion of the total tumor population (TP) represented by (TS_H):

   $$Tumor\ Population\ Fraction\ (TPFx) = TS_H \diagup TP$$

3) Represents a minimal portion of the total normal population as determined by the complement of the total normal population (NP_H) represented by the number of normal samples (NS_H) present in CP_H.

   $$Complement\ Normal\ Population\ Fraction\ (CNPFx) = [1 - (NS_H \diagup NP_H)]$$

4) Exhibits a significantly higher expression value than the low expression cluster as indicated by the difference between the unlogged medoid expression value of the high expression cluster (EV_H) and the of the low expression cluster (EV_L)

   $$Expression\ Difference\ (ED) = \alpha log(EV_H) - \alpha log\ (EV_L)$$

The priority ranking of each transcript candidate is then determined by:

$$Transcript\ Rank\ = TFx * TPFx * CNPFx * ED$$

The ranked novel assemblies now undergo *translation potential assessment* (**Fig. 2**) to elucidate those sequences possessing the best potential for translation into unique peptide constructs. The coding sequence of each transcript is translated in all three frames using the EMBOSS *transeq* tool [15,16], and the longest open reading frame (ORF) is selected. This sequence is aligned to the peptide sequences for all transcripts of the hg19 reference gene most closely related to the novel isoform using EMBL-EBI Clustal Omega [17] and EMBOSS *showalign* [15]. Any candidate lacking a start site shared by at least one of the reference sequences or possessing an ORF identical to any of the reference transcripts is removed from the dataset. Of the remaining candidates, the unique portion(s) of each transcript are aligned to the entire hg19 reference genome using BLAT [18], and any sequence(s) found to align elsewhere in the genome are also discarded. As longer candidate epitope sequences provide more opportunity for a true immunologic target, the remaining transcripts undergo *epitope candidate ranking* (**Fig. 2)** to take the length of the potential epitope into account.

$$Epitope\ Rank = Transcript\ Rank * Candidate\ Epitope\ Length$$

Twenty of the top-ranked known transcripts (**Table 3A)** and twenty of the top-ranked novel epitopes (**Table 3B**) are listed in **Table 3**. A number of the known transcripts provided in **Table 3A** are already known to be associated with breast cancer. The miR492 and miR622 micro-RNAs are found to have expression signatures correlated with specific breast cancer subtypes [19], and miR492 is particular is associated with supporting hepatic cancer progression through targeting of PTEN [20]. The cellular retinoic acid binding protein (CRABP2) is jointly regulated with estrogen receptor alpha and retinoic acid receptor alpha in human breast cancer cells [21]. The guanine nucleotide-binding protein subunit beta-2-like 1 (GNB2L1, or RACK1) has been reported as a predictor for pool clinical outcome in breast cancer patients and has potential to be an independent biomarker for diagnosis and prognosis of breast cancer. Upregulation of S100A11 is reported in a variety of metastatic cancers and is essential for the efficient repair of the plasma membrane and for the survival of highly motile cancer cells [22], while overexpression of S100A14 modulates HER2 signaling in breast cancer [23]. Interferon alpha-inducible protein 6 (IFI6, or G1P3) promotes hyperplasia, tamoxifen resistance, and poor patient outcomes in breast cancer [24]. The estrogen-responsive anterior gradient 2 (AGR2) influences dissemination of metastatic breast cancer cells and may be useful as a marker in identification of circulating tumor and metastatic cells in sentinel lymph nodes. It is also a promising drug target and prognostic indicator [25].

The prevalence of breast cancer associated genes residing in high-ranking positions of this dataset lends significant support to the functionality of our pipeline as well as validity to the top-ranking epitope candidate results (**Table 3B**). In fact, even amongst the top-ranked epitope candidates shown in **Table 3B**, there are a number of cancer-related genes, including thymosin beta-10 (TMSB10, G-actin sequestration and breast

A

Gene	Low Expr (EV_L)	High Expr (EV_H)	Tumor Fxn (TF_X)	Tumor Popn Fxn (TPF_X)	Nrml Popn Fxn ($CNPF_X$)	Transcript Rank
TMSB10P1	9.71	12.36	0.85	0.39	0.13	1266.77
MIR492	0.00	11.45	0.74	0.99	0.66	698.78
RPL10	0.03	10.03	0.85	0.62	0.22	432.94
B2M	9.76	11.27	0.72	0.57	0.44	368.21
PABPC1	0.00	9.00	0.92	0.58	0.10	245.02
RPLP1	0.00	10.70	0.71	0.98	0.80	231.37
RPS24	0.00	9.26	0.80	0.72	0.36	226.93
CRABP2	1.80	8.45	0.97	0.69	0.04	219.88
GNB2L1	0.00	9.48	0.74	0.58	0.39	188.28
TFF1	0.15	8.98	0.93	0.42	0.06	185.09
RPL30	0.00	9.01	0.87	0.48	0.14	183.29
MYL6P1	8.04	10.22	0.71	0.41	0.33	179.36
RPL30	0.00	10.46	0.89	0.15	0.03	176.55
S100A11	5.79	9.25	0.76	0.98	0.59	168.61
MIR622	0.00	8.22	0.88	0.72	0.18	153.28
NPM1	0.00	8.70	0.80	0.70	0.34	152.20
S100A14	0.00	8.02	0.94	0.65	0.08	145.67
RPLP0	0.00	8.50	0.80	0.81	0.40	139.27
IFI6	6.06	8.69	0.92	0.47	0.08	136.30
AGR2	1.55	8.30	0.81	0.82	0.37	130.77

B

Gene	Low Expr (EV_L)	High Expr (EV_H)	Tumor Fxn (TF_X)	Tumor Popn Fxn (TPF_X)	1-Nrml Popn Fxn ($CNPF_X$)	Transcript Rank	Epitope Length	Epitope Rank
TMSB10	11.25	12.74	0.86	0.82	0.26	2303.20	15	34548.03
KRT18	0.98	8.95	0.91	0.66	0.12	261.29	32	8361.19
ANXA2	7.07	10.22	0.77	0.96	0.58	325.58	12	3906.90
SEC61A1	0.00	6.70	0.80	0.57	0.28	34.06	100	3406.22
COL1A1	0.00	8.39	0.94	0.59	0.08	169.86	20	3397.28
MUC1	0.00	5.78	0.86	0.55	0.17	20.93	141	2951.29
SPINT2	2.61	7.20	0.77	1.00	0.59	44.53	61	2716.34
IGKV3-20	0.00	8.59	0.73	0.50	0.36	90.12	21	1892.46
SEC61A1	0.28	6.05	0.80	0.46	0.23	18.27	100	1827.15
TPD52	0.81	4.88	0.89	0.80	0.19	16.21	93	1507.84
GATA3	0.00	5.76	0.95	0.74	0.08	34.48	43	1482.68
TMED2	0.20	5.57	0.94	0.69	0.09	27.52	53	1458.58
HDLBP	4.41	6.51	0.81	0.46	0.21	20.70	67	1386.75
COL8A2	0.56	4.35	0.95	0.43	0.05	7.24	163	1179.65
HM13	2.00	5.85	0.94	0.57	0.08	26.59	44	1169.84
DDX23	0.01	4.74	0.73	0.55	0.41	6.07	176	1067.60
UGT2B11	0.00	7.40	0.87	0.08	0.03	11.66	86	1002.55
HNRNPM	0.00	5.43	0.85	0.56	0.19	16.13	62	1000.24
GTF2H5	0.00	8.27	0.77	0.69	0.41	96.30	10	963.00
LAMB2	1.88	5.14	0.60	0.40	0.53	3.52	243	854.16

Table 3. Twenty of the top ranked known (A) and novel (B) transcript candidates predicted by the epitope discovery pipeline in terms of 'transcript rank' for known transcripts and 'epitope rank' for predicted epitope sequence of novel isoforms.

cancer cell motility) [26], keratin 18 (KRT18, tumor dedifferentiation and loss of estrogen and progesterone receptors) [27], and annexin A2 (ANXA2, invasion augmentation of multidrug-resistant breast tumor cells) [28].

We believe the modifications made to the analytical pipeline improve its efficiency in predicting tumor-specific transcripts and neoantigen candidates. The ranking system helps to prioritize those transcripts and neoantigen candidates most suitable for future research as immunological targets, and the automated assessment of translation potential and isolation of novel sequences has reduced computational time from hours to minutes. This computational procedure is intended for use with RNAseq data obtained from enrolled patient tumors to verify our results.

Identify small molecule agents enhancing tumor cell apoptosis and CTL killing [Task 12]

As outlined in Aim 4 of the proposal, clinical efficacy of T cell-based therapies will be enhanced in combination with agents promoting tumor cell apoptosis. Support for this idea recently has been published showing

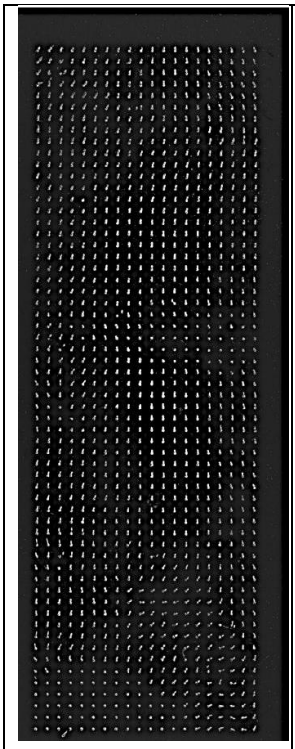

Figure 3. PBMC attachment to Con A spotted slide.

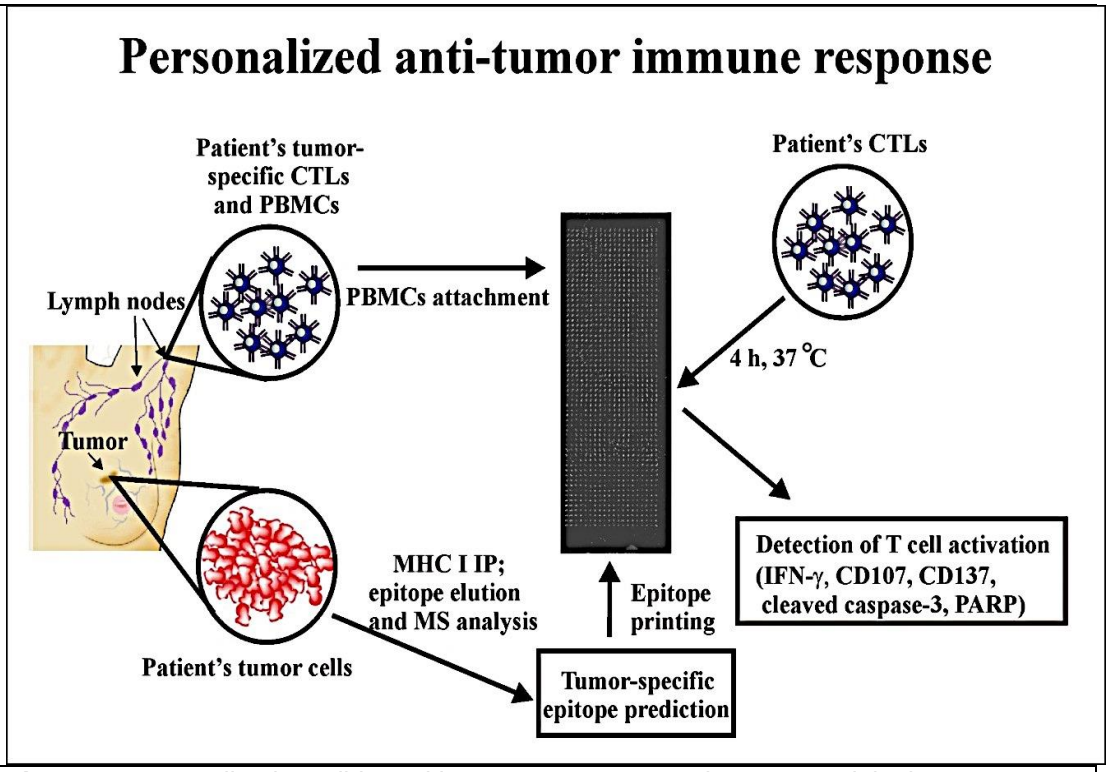

Figure 4. Personalized T cell-based immune response. IP, immunoprecipitation; MS, mass spectrometry

chemotherapy can synergize with CTL-mediated killing [29]; however, chemotherapeutic agents can also inhibit T cell function.

We are continuing our work in this area to identify drugs nontoxic to normal cells by developing T cell cytotoxicity assays using the peptides we previously characterized (see Task 5 above). We have determined the printing of Con A to specific spots on a slide allows attachment of PBMCs to these spots (**Fig. 3)**. We also show attachment to Con A does not affect cell functionality, such as cell proliferation rate and ability of siRNA to reduce gene expression.

Finally, we designed a protocol to personalize T cell-based treatment (**Fig. 4)**. In this protocol, we will perform MHC I immunoprecipitation and epitope elution from patient tumor tissue, as we did with the breast carcinoma cell lines, followed by mass spectrometry analysis. Tumor-specific epitopes will be selected by gene expression analysis of the corresponding proteins. These epitopes, altogether with Con A, will be printed on a slide. Next, we will extract PBMCs and cytotoxic T lymphocytes (CTLs) from the same patient and allow the PBMCs to bind to the Con A spotted slide. Because the PBMCs are from the same patient, we do not need to know the type of MHC I alleles present in the patient. The slide, with attached PBMCs, will then be incubated with the patient's CTLs, and T cell activation will be detected using IFN-γ, CD137, CD107, and other T cell activation markers. We plan to test this protocol using tumors from breast cancer patients consented to the project by the City of Hope working group.

KEY RESEARCH ACCOMPLISHMENTS:

- Determined which of the 170 MHC I-loaded eluted epitopes identified in the previous funded year exhibited the ability to activate T cells. Eleven sequences were characterized as immunogenic, several of which were found in multiple cell lines.
- Modified open source MiTCR software to allow matching sequence reads from the alpha and beta chains of a single TCR followed by calculation of clonotype frequencies. Input to the program is raw sequence data from single TCRs generated by the Slansky team. The software is repackaged as CompleteClone.
- Modified the epitope discovery pipeline developed in the previous funding year for *in silico* prediction of breast cancer epitopes from RNAseq data. A more robust method for transcript and neoantigen candidate

prioritization was instituted, and an automated approach for validating transcription potential of novel isoforms and isolation of potential neoantigen sequences was developed.

- Designing a protocol for personalization of T cell-based therapy through direct observation of tumor-derived T cell activation against epitopes eluted from the same patient.

CONCLUSION:

The focus of the Spellman/Gray work group over the past year has been upon the generation of materials, tools, and data for the purpose of aiding and supporting the research and findings of the entire multi-team collaboration endeavoring to identify antigenic targets for breast cancer-infiltrating T cells. We have identified a number of candidates in breast cancer tissues as well as breast cancer cell lines, utilizing a variety of analytical methods. The epitope discovery pipeline is proof of concept of *in silico* epitope discovery from RNAseq data. It aids in the definition of the protein-epitope relationship by enlarging the knowledge base of protein-encoding transcripts beyond the protein models existing in public databases and by restricting the analyses to only the expressed transcripts. The results produced by this pipeline along with the MHC-I-bound epitopes identified by mass spectrometry in breast cancer cell lines will be used to rank epitopes for further characterization and development as therapeutic targets.

PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS:

No publications, abstracts, or presentations to report.

INVENTIONS, PATENTS, AND LICENSES:

No inventions, patents, or licenses to report.

REPORTABLE OUTCOMES:

NBCC/Artemis Project: We have developed a computational pipeline, coined CompleteClone, which analyzes raw TCR sequence data from single T cells, independently identifies the CDR3 sequence and VDJ alleles of the alpha and beta chains, matches the alpha and beta reads for individual TCR clonotypes, and calculates clonotype frequencies for the T cell clone. The software is currently used only with sequence data produced by the Slansky team following their single-cell emulsion RT-PCR technique; however, it can be packaged and shared for use with others for similar purposes.

OTHER ACHIEVEMENTS:

No other achievements to report.

REFERENCES:

1. Zajac AJ, Murali-Krishna K, Blattman JN, Ahmed R (1998) Therapeutic vaccination against chronic viral infection: the importance of cooperation between CD4+ and CD8+ T cells. Curr Opin Immunol 10: 444–449.
2. cBioPortal for Cancer Genomics (n.d.). Available: http://www.cbioportal.org/public-portal/. Accessed 16 October 2014.
3. Home - The Cancer Genome Atlas - Cancer Genome - TCGA (n.d.). Available: http://cancergenome.nih.gov/. Accessed 26 March 2013.
4. Kapushesky M, Emam I, Holloway E, Kurnosov P, Zorin A, et al. (2009) Gene Expression Atlas at the European Bioinformatics Institute. Nucleic Acids Research 38: D690–D698. doi:10.1093/nar/gkp936.
5. Edgar R (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research 30: 207–210. doi:10.1093/nar/30.1.207.
6. Activation-induced expression of CD137 permits detection, isolation, and expansion of the full repertoire of CD8+ T cells responding to antigen without requiring knowledge of epitope specificities (n.d.). Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1896114/?report=classic. Accessed 16 October 2014.
7. In vitro methods for generating CD8+ T-cell clones for immunotherapy from the naïve repertoire (n.d.). Available: http://www.sciencedirect.com/science/article/pii/S0022175905004291. Accessed 16 October 2014.
8. Bolotin DA, Shugay M, Mamedov IZ, Putintseva EV, Turchaninova MA, et al. (2013) MiTCR: software for T-cell receptor sequencing data analysis. Nat Meth 10: 813–814. doi:10.1038/nmeth.2555.

9. R Core Team (2014) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Available: http://www.R-project.org/.

10. Java SE - Downloads | Oracle Technology Network | Oracle (n.d.). Available: http://www.oracle.com/technetwork/java/javase/downloads/index.html. Accessed 6 May 2014.

11. Wickham, Hadley (2011) The Split-Apply-Combine Strategy of Data Analysis. J Stat Softw 40: 1–29.

12. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25. doi:10.1186/gb-2009-10-3-r25.

13. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25: 1105–1111. doi:10.1093/bioinformatics/btp120.

14. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnology 28: 511–515. doi:10.1038/nbt.1621.

15. Rice P, Longden I, Bleasby A (2000) EMBOSS: The European Molecular Biology Open Software Suite. Trends in Genetics 16: 276–277. doi:10.1016/S0168-9525(00)02024-2.

16. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, et al. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 38: W695–W699. doi:10.1093/nar/gkq313.

17. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology 7. Available: http://www.nature.com/doifinder/10.1038/msb.2011.75. Accessed 17 September 2013.

18. Kent WJ (2002) BLAT—The BLAST-Like Alignment Tool. Genome Res 12: 656–664. doi:10.1101/gr.229202.

19. Riaz M, van Jaarsveld MT, Hollestelle A, Prager-van der Smissen WJ, Heine AA, et al. (2013) miRNA expression profiling of 51 human breast cancer cell lines reveals subtype and driver mutation-specific miRNAs. Breast Cancer Research 15: R33. doi:10.1186/bcr3415.

20. Jiang J, Zhang Y, Yu C, Li Z, Pan Y, et al. (2014) MicroRNA-492 expression promotes the progression of hepatic cancer by targeting PTEN. Cancer Cell International 14. Available: http://www.cancerci.com/content/14/1/95. Accessed 20 October 2014.

21. Lu M, Mira-y-Lopez R, Nakajo S, Nakaya K, Jing Y (2005) Expression of estrogen receptor alpha, retinoic acid receptor alpha and cellular retinoic acid binding protein II genes is coordinately regulated in human breast cancer cells. Oncogene 24: 4362–4369. doi:10.1038/sj.onc.1208661.

22. Jaiswal JK, Lauritzen SP, Scheffer L, Sakaguchi M, Bunkenborg J, et al. (2014) S100A11 is required for efficient plasma membrane repair and survival of invasive cancer cells. Nature Communications 5. Available: http://www.nature.com/doifinder/10.1038/ncomms4795. Accessed 20 October 2014.

23. Xu C, Chen H, Wang X, Gao J, Che Y, et al. (2014) S100A14, a Member of the EF-hand Calcium-binding Proteins, Is Overexpressed in Breast Cancer and Acts as a Modulator of HER2 Signaling. Journal of Biological Chemistry 289: 827–837. doi:10.1074/jbc.M113.469718.

24. Cheriyath V, Kuhns MA, Jacobs BS, Evangelista P, Elson P, et al. (2012) G1P3, an interferon- and estrogen-induced survival protein contributes to hyperplasia, tamoxifen resistance and poor outcomes in breast cancer. Oncogene 31: 2222–2236. doi:10.1038/onc.2011.393.

25. Salmans ML, Zhao F, Andersen B (2013) The estrogen-regulated anterior gradient 2 (AGR2) protein in breast cancer: a potential drug target and biomarker. Breast Cancer Res 15: 204. doi:10.1186/bcr3408.

26. Mælan AE, Rasmussen TK, Larsson L-I (2006) Localization of thymosin β10 in breast cancer cells: relationship to actin cytoskeletal remodeling and cell motility. Histochemistry and Cell Biology 127: 109–113. doi:10.1007/s00418-006-0208-z.

27. Seon-Ah H, Soo LY, Kee KH, Jinah Y, Sanghee K, et al. (2011) The prognostic potential of keratin 18 in breast cancer associated with tumor dedifferentiation, and the loss of estrogen and progesterone receptors. Cancer Biomarkers: 219–231. doi:10.3233/CBM-2012-0250.

28. Zhang F, Zhang L, Zhang B, Wei X, Yang Y, et al. (2009) Anxa2 plays a critical role in enhanced invasiveness of the multidrug resistant human breast cancer cells. J Proteome Res 8: 5041–5047. doi:10.1021/pr900461c.

29. Ramakrishnan R, Assudani D, Nagaraj S, Hunter T, Cho H-I, et al. (2010) Chemotherapy enhances tumor cell susceptibility to CTL-mediated killing during cancer immunotherapy in mice. J Clin Invest 120: 1111–1124. doi:10.1172/JCI40269.

APPENDICES:

No appendices to report.